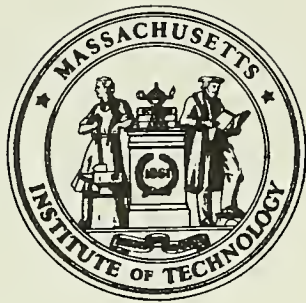




LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY



Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

<http://www.archive.org/details/onrelationsbetwe00fish>

**working paper
department
of economics**

ON THE RELATIONS BETWEEN ECONOMETRIC
SECTORAL AND MACRO MODELS

by

Franklin M. Fisher

Number 215

February 1978

**massachusetts
institute of
technology**

**50 memorial drive
cambridge, mass. 02139**

ON THE RELATIONS BETWEEN ECONOMETRIC
SECTORAL AND MACRO MODELS

by

Franklin M. Fisher

Number 215

February 1978

Dewey



ON THE RELATIONS BETWEEN ECONOMETRIC

SECTORAL AND MACRO MODELS^{*}

Franklin M. Fisher,
Massachusetts Institute of Technology

1. Introduction

This paper is concerned with the methodological and practical problems which arise in the relationships between two types of econometric model -- macro models, which deal with the entire economy, usually on a fairly aggregate basis, and sectoral models, which deal with portions of the economy, usually on a somewhat more detailed basis. Regardless of which type of model is one's primary concern in a given case, it is easy to see that some thought may have to be given to the existence of the other type.

Thus, for example, consider the point of view of an investigator modelling the economic relationships of a particular sector -- say, the demand and supply functions of the market for a particular commodity. Almost inevitably such a model will contain variables determined outside the sector of interest but endogenous to a model of the economy as a whole. How should such variables be treated in the estimation process? Further, how should they be treated for purposes of forecasting or simulation?

Very often (although certainly not always), such outside-determined variables will be aggregate ones. Some measure of aggregate income, for

^{*} Paper to be presented at The School of Economic Planning for Government and Industry, SOGESTA, Urbino, Italy, January 1978.

example, commonly appears in demand functions. (However, so, too, do the prices of other commodities, which are not aggregates.) If one is lucky, existing macro models will provide adequate forecasts of such magnitudes. If one is only as fortunate as most econometricians, however, such forecasts, however adequate on their own grounds, will be of variables that differ at least slightly in their definition or data sources from the ones needed in the sectoral model. A sub-case, as it were, of the question asked above as to forecasting and simulation can thus be put as follows: I have available to me forecasts and data from existing macro models. How can such forecasts be best adapted for sectoral modelling purposes? To put it another way, to what extent should I build my sectoral model to conform to the protocols, so to speak -- the definitions and usages of an existing macro model? Must this be done at the outset, or is it better to wait until the forecasting and simulation phase of the work and adapt then one model or the other? Indeed, given that the sector being modelled is properly thought of as embedded in the larger economy modelled in a macro model, is it not the case that proper modelling technique requires that such embedding be made explicit and both models be built and estimated together? To what extent, if any, can such a complex effort be avoided?

I have so far spoken from the point of view of the investigator whose primary interest is in a particular sector. Somewhat related problems arise, however, when considering the point of view of the investigator who is modelling the entire economy. If such a model is highly detailed, it is likely to be large enough to require work on different parts by different people. Even where this does not happen at the outset because

of the aggregative nature of the model, such aggregation is never more than a possibly useful approximation. The macro modeller must realize that, either actually or potentially, the broad outline of his model will be fleshed out and the detail articulated. To what extent ought he to take this into account in planning the model? To put it another way, if parts of the model (sectors) are to be done in detail by others -- either associates in the initial model building effort or later investigators interested in particular sectors -- to what extent can such efforts be decentralized? When must such labors be coordinated with the central model building effort? And when must the results of such articulated sectoral modelling be taken into account?

The present paper offers my thoughts on these and related questions. While I have little, if anything, to say about them that I or others have not said before, there is perhaps some point to drawing such discussions together and dealing with the relations between sectoral and macro models in a moderately systematic way.

2. What is a Sector?

It is clear from consideration of the questions raised above that both sectoral and macro models are best thought of as coming from an underlying ideal model of the economic system which has many equations

and considerable detail. Such an ideal model being only ideal, both sectoral and macro models abstract from it in different ways. Thus, macro models compress such an ideal model by aggregating and ignoring much of the detail. Sectoral models, on the other hand, truncate the underlying model by dealing with only some of its relationships and treating many of its variables as somehow determined outside the truncated model being investigated.¹

While one might consider any such truncation to be a sectoral model, it is more useful to be a little more restrictive and to ask what characteristics such truncations are likely to have that would lead us to consider them to be interesting sectors. One might, for example, consider the equation for the demand for pigs and that for the supply of money, but most economists would not consider the two together to be the equations of a sector (although where monetary theory is concerned, one can never be sure). Those two equations either put together too much or else not enough to be a sector.

What then do we mean (or, perhaps, should we mean) by a "sector"? The answer is not unrelated to the analysis which follows, or to the questions already posed. This is so because there is not much point in this whole discussion if it is going to turn out that sectors generally can not be thought of as separable or nearly separable entities. Hence, I shall build some of the analysis into the definition and elaborate on

¹ Depending on the level of detail which one considers the underlying model to have, sectoral models aggregate also. Certainly this is true if one thinks of the underlying model as operating at the level of individual households and firms. For our purposes, however, this is not particularly helpful, and we may as well think of the levels of aggregation of the underlying model and the sectoral models derived from it as being the same.

it later, rather than take a very broad definition and then show the problems that arise. This is, of course, only a matter of language. I shall later discuss the issues involved in giving sectoral treatment to subsets of equations which do not meet the definition.

Accordingly, I propose the following as a sensible working definition. From the equations of the underlying model, choose a certain subset. Corresponding to that subset of equations will be a subset of current endogenous variables (equal in number) which can be thought of as determined by the chosen equations, given the values of the remaining variables (some of which can be endogenous to the full model). The subset of equations (and the corresponding subset of endogenous variables) is properly considered a sector if either of two conditions hold.

The first such condition is that there is no feedback (either simultaneous or lagged) in the underlying model from the variables of the proposed sector to the remaining endogenous variables of the larger model. This is the case in which forces in the larger model determine the variables in the sector, but not vice versa.

In general, this condition can not be expected to hold. In the general case, there are such feedback effects, and one ignores them in treating the subset of equations chosen as determining the corresponding endogenous variables. Accordingly, the alternative condition for considering the subset as a sector is that the importance of the feedback effects which must be ignored in this way is small.

An example may help to clarify matters. Suppose that one is interested in studying the market for milk. Consider equations for the demand and supply of milk as the equations of a proposed sector, with corresponding endogenous variables being the price and quantity of milk.

In general, one would expect income to enter the demand function for milk.

This means that there is a feedback effect to the larger model because the price and quantity of milk affect the income of dairy farmers. In most economies, however, we would consider the milk equations to comprise a sensible sector because the importance of this effect on total income is small. In such a situation, it makes sense to study milk supply and demand taking income to be exogenous as a first approximation.¹ Note, however, that if we were looking at a model of milk supply and demand in a community with a high proportion of dairy farmers, this would not be the case. In such a situation, one would go badly astray in estimation, and also in simulation and forecasting, if one decided provisionally to ignore such feedback effects on income.

It is because one can go badly astray when ignoring large feedback loops that I have chosen to define a "sector" in the way I have. Obviously, definitions are only that and it might be said that the definition in this paper departs to some extent from common usage. One does, for example, talk of the "energy sector" even though there are large feedbacks between energy and other parts of the economy. However, ignoring those feedbacks makes it difficult to do anything with equations for the energy industry taken in isolation and it seems to me better to say that what one has is only part of a "sector." Clearly, this has nothing to do with the analysis below. Further, the appropriateness of ignoring feedback loops may depend upon the purposes to which the model is going to be put. In any case, the consequences of ignoring feedback loops are discussed below to some extent.

Diagrammatically, the existence of a sector is depicted in Figure 2.1. In this diagram, arrows denote causal effects (either current or lagged). The sector is depicted as having three equations and one internal feedback loop (I shall have more to say about this in a moment). The variables of the rest of the model influence the variables of the sector through the arrow denoted "a", and are themselves influenced by the latter variables through the arrow denoted "b". For the sector to be a sensible one, it has to make sense to ignore at least one of these two arrows so that the entire feedback effect between the sector and the rest of the model is small.

Note that it will do if either a or b can be ignored. In the milk example just given, the effect of income on the demand for milk can be represented by an arrow like a. What is to be ignored is the effect of

¹The extent to which this makes sense, however, depends in part on what one is trying to do. See the extended discussion in §4 below.

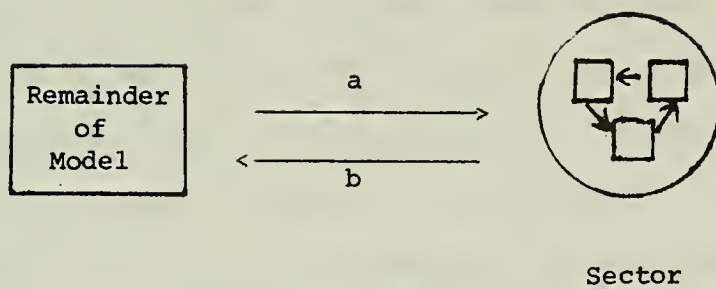


Figure 2.1
Diagrammatic Representation of a Sector

the milk market variables on income, an effect such as $\underline{b}$. But of course it would be enough if it were the other way around. This would be the case in which there was a strong effect of milk market variables on income, but in which income did not strongly affect the demand for milk. Also in such a case, it would make sense to deal separately with the milk market as a first approximation and then go on to consider its effects on the rest of the economy.

If this particular example seems implausible looked at in this way, consider the following. There is a definite symmetry between the way in which one treats a sector and the way in which one treats the rest of the model. Thus, continuing the milk example, and supposing it to be that pictured in the diagram, we have agreed that it makes sense to treat the milk market as a separate sector if the influences denoted by $\underline{b}$ (the effects of the milk market variables on income) are small. If this is the case, however, then one can also treat the remainder of the model separately, that is, without worrying about its effects on the milk market. Those effects, even though strong, can be analyzed later because they do not come back in a significant way into the remainder of the model. But this means that the remainder of the model can itself be considered a sector either because the arrows pointing out of it are small, or because the arrows pointing into it are small. In either case, the feedback effects will be small and (also in either case) the possibility of the treatment of the subset of equations as a sector implies (and is implied by) the possibility of the similar treatment of the complement of that subset in the equations of the model.

There are a few points in the above discussion that require some

further elaboration. The first of these is the matter of lagged versus simultaneous effects. I have lumped both of these together and have spoken of feedback loops as involving either one. In fact, however, it may make some difference which is involved. If the feedback effects between a sector and the rest of the model only operate with a lag, then it is not hard to see that for estimation purposes, the sector may be a sensible one, even though for purposes of simulation and forecasting it is not. This will be so because the variables of the remainder of the model can be taken as predetermined in the estimation of the equations of the sector, provided appropriate assumptions are made about the disturbances. Those assumptions, however, may be hard to justify except as a matter of convenience, and in practice, the distinction here is much fuzzier than might at first be supposed. I have therefore not drawn it in the general definition given, although I shall have to return to these matters below.

Second, the above discussion assumes that there is a natural association between endogenous variables on the one hand and equations on the other, so that the choice of which equations are to be treated as a sector implies the choice of a corresponding number of endogenous variables which can be thought of as somehow determined by those equations. In effect, this amounts to assuming that each equation has a natural normalization rule which specifies a particular left-hand side endogenous variable which it determines in some sense, given the variables on the right-hand side. Yet, clearly the endogenous variables of a full simultaneous model are all determined jointly, given the predetermined variables and the disturbances by all the equations of the model. This is what simultaneity is all about. Is this not a contradiction?

I have had a good deal to say about this issue in other contexts,¹ and, since it is only loosely related to the principal concerns of this paper, I shall give only a part of that discussion in full here. There are at least two (related) ways of seeing that there is no real problem. The first of these is rather closely related to the matters we are discussing. The equations of a simultaneous model are not formulated as an organic whole; rather, they are formulated one at a time and then treated together for purposes of estimation, simulation, and forecasting. To put it a little more concretely, the consumption function, for example, has a theoretical life of its own. One can imagine the same consumption function appearing in several different models, each of which has a different investment function. Indeed, one can imagine the same consumption function appearing in a model of an economy in which income and all variables except consumption have been somehow set by government fiat. The behavior described by that consumption function is the determination of consumption decisions, given other variables, and it is natural, as the name suggests, to treat consumption as the endogenous variable appearing on the left-hand side, whereas it would not be natural to normalize that equation and treat it as in any serious sense determining income.

Note the relation to the discussion above. In terms of the milk market example, it is natural to think of income as influencing the price and quantity of milk through the demand equation for milk. The effect of the milk market variables on income may or may not be important, but it occurs in quite another place in the model. In considering whether or not to ignore that effect, there is no doubt about which equations are

¹Fisher (1965), (1970a,b).

involved. One would never think of the demand for milk as determining income; the gedanken experiment corresponding to normalizing the demand for milk clearly involves thinking of the demand equation as embedded in a world in which income -- not milk demand -- is determined from outside. Were this not so, no treatment of sectors would ever be possible.

Alternatively, one can think of such natural normalization rules as arising from the approximate nature of simultaneous equations. If one believes that simultaneity is only an approximation and that the effects described in the model in fact take small amounts of time to operate, then each equation in the model is an approximation to an underlying equation with only one unlagged variable. That variable can be put on the left-hand side and is the one which, even in the simultaneous approximation, can be thought of as naturally associated with the equation.

In practice, it isn't always as straightforward as this. Normalization rules may simply not be clear. In the supply of milk equation, for example, is it price or quantity which is the natural left-hand side variable? I think it is preferable to think of such cases as ones in which the normalization rules are unknown, or in which some of the equations of the model have been substituted out, as opposed to being cases in which such rules do not exist. The answer to the question about milk supply, for example, depends on a close examination of the pricing behavior being modelled, and may require the addition of an explicit pricing equation. In any case, such ambiguities seem to me invariably to occur within sets of equations which can be treated as sectors, rather

than going outside of sectoral lines (in the milk case, for example, there isn't any doubt as to the fact that it's the quantity and price of milk that are associated with the milk sector equations even though there may be some doubt as to which variable should be thought of as going with which equation). I shall henceforth assume that these matters present no difficulties.

The final point to be discussed before proceeding is rather simpler. In the diagram, I have included a within-sector feedback loop. This is not necessary for the definition of a sector, but I shall generally suppose there to be one anyway. If there is not such a loop, the sector under consideration can, for formal purposes, be broken down further into smaller sectors and each one separately analyzed. There is, of course, nothing to prevent a sector from consisting of a single equation, and I do not mean to rule out such a case formally. However, the issues presented by such a case are just a subclass of those presented by cases which do have feedback loops and I shall thus generally be thinking in terms of the latter. Hence, I shall generally ignore the possibility of breaking a sector into subsectors and shall suppose all the equations of a sector to be connected through a feedback loop, even though this may not always be the case in actual practice.

3. Block Recursiveness: Identification and Estimation

In order to consider the problems which arise in the estimation of sectoral models, it is useful to give a somewhat more formal description of what is going on than I have so far done. Accordingly, suppose that after normalization of each equation, the equations of the underlying model can be written as:

$$(3.1) \quad y = By + \Gamma z + u$$

where y is an M -component vector of endogenous variables; z is a Λ -component vector of predetermined variables; u is an M -component vector of disturbances; B is an $M \times M$ matrix of coefficients; and Γ is an $M \times \Lambda$ matrix of coefficients. The components of B and Γ are to be estimated. For the normalization rules to make sense, B must have zeroes everywhere on its principal diagonal.

I have not been explicit about lagged endogenous variables in (3.1) because such an explicit treatment is not necessary at this stage. For estimation purposes, the principal issue involving lagged endogenous variables is that of whether they can be treated as predetermined. Since this is also a principal issue in the treatment of endogenous variables not included in the sector to be estimated, we can for many purposes treat the two problems together. To do this, it will sometimes be convenient to think of the equations of the model as including not only

the equations which are usually thought of as structural equations, but also the lagged versions thereof. If we do so, then lagged endogenous variables just become additional current endogenous variables included in the vector, y . It is true that extending the model in such a way involves a potentially infinite number of equations and variables, but, at the expository level at which I shall work, it is convenient just to ignore this. To put it another way, the problems involved with lagged endogenous variables are sufficiently similar to those involved anyway that we can analyze the sectoral estimation problem in a model without such variables and then interpret the results so as to include them.

Now suppose that the equations and the corresponding endogenous variables are renumbered (if necessary) and then grouped into subsets. Thus, partition y , B , Γ , and u as:

$$(3.2) \quad y = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^C \end{pmatrix}; \quad B = \begin{pmatrix} B^{11} & B^{12} & \dots & B^{1C} \\ B^{21} & B^{22} & \dots & B^{2C} \\ \vdots & \vdots & & \vdots \\ B^{C1} & B^{C2} & \dots & B^{CC} \end{pmatrix};$$

$$\Gamma = \begin{pmatrix} \Gamma^1 \\ \Gamma^2 \\ \vdots \\ \Gamma^C \end{pmatrix}; \quad u = \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^C \end{pmatrix},$$

where y^k denotes the subvector of y corresponding to the k^{th} group of endogenous variables and the partitioning of B , Γ , and u corresponds to this. Let Σ be the contemporaneous variance-covariance matrix of the disturbances and partition it also in the obvious way.

A square matrix, partitioned as is B or Σ , with square submatrices (blocks) on the principal diagonal, is called block-triangular if all blocks above that diagonal are zero. In the case of B this would mean that $B^{hk} = 0$ for all $k > h$. Such a matrix is called block-diagonal if all blocks both above and below the principal diagonal are zero. For Σ , for example, this would mean $\Sigma^{hk} = 0$ for all $h \neq k$.

Definition 3.1: A simultaneous equation system such as (3.1) is called block-recursive¹ if there exists a partitioning such as (3.2) in which B is block-triangular and Σ is block-diagonal.

It is not hard to see that a block-recursive system is one in which each group of variables and equations can be treated separately. Indeed, the model breaks up very neatly into sectors. To see this, we begin by writing out the sets of equations in (3.1) using the partitioning (3.2) and the block-triangularity of B .

$$\begin{aligned}
 y^1 &= B^{11}y^1 + \Gamma^1z + u^1 \\
 y^2 &= B^{21}y^1 + B^{22}y^2 + \Gamma^2z + u^2 \\
 &\vdots \\
 y^C &= B^{C1}y^1 + B^{C2}y^2 + \dots + B^{CC}y^C + \Gamma^Cz + u^C.
 \end{aligned}
 \tag{3.3}$$

The first set of equations here state that the variables in y^1 are simultaneously determined given only the predetermined variables and the first

¹This terminology was introduced in Fisher (1961). If each block has only one equation, the model is said to be recursive. See Wold (1953).

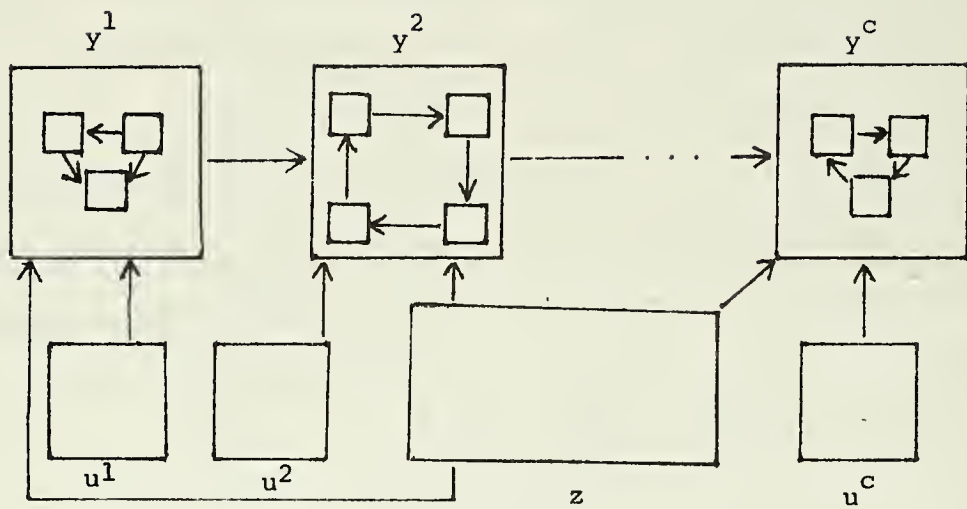


Figure 3.1
A Block-Recursive Model

set of disturbances. They do not depend on the endogenous variables in higher numbered sets. Similarly, the endogenous variables in y^2 are simultaneously determined given only the predetermined variables, the corresponding set of disturbances, and the endogenous variables from lower-numbered sets.

In terms of a causal flow diagram such as Figure 2.1 above, the situation is depicted in Figure 3.1. The point is that while there can be feedbacks within each group of equations, there are no feedbacks operating between groups. This is precisely the situation, discussed in the previous section, which leads us to treat each set of equations (and the corresponding set of endogenous variables) as a sector.

To what extent does such treatment carry over into estimation? It is clearly tempting to estimate the equations of each sector separately, treating as predetermined the endogenous variables from lower-numbered sectors and ignoring those from higher-numbered sectors altogether. Does this work?

Consider first the equations of the first sector, those corresponding to y^1 . Such equations contain no endogenous variables other than y^1 , so it is clearly safe to estimate them as they stand. The fact that y^1 appears again later down the line is irrelevant; the equations of the first sector are in the form of a classical simultaneous equations system all by themselves. Indeed, it is not hard to show¹ that, so far as identification is concerned, an equation in the first sector is identifiable in the full model if it is identifiable considering the equations of the first sector only; the block-triangularity of B implies that any linear

¹See Fisher (1966) for this and similar later statements about identification.

combination of equations which gives non-zero weight to equations of later sectors must involve non-zero coefficients for endogenous variables not in the first sector. (B is block-triangular, and, we assume, non-singular, whence every B^{kk} is also non-singular.) Essentially the same thing applies to estimation.

Now how about the equations of the second sector, those corresponding to y^2 ? At first glance, it looks as though the same argument will apply. Those equations contain, in addition to y^2 , only the variables of z and of y^1 . But the variables of z are predetermined and present no difficulty, while, as we have just observed, the variables of y^1 depend only on the variables of z and not at all on the variables of y^2 . Hence it looks as though we can treat y^1 as predetermined when dealing with the second (or higher-numbered) sectors.

There is a big catch in this argument, however, and it can be seen in more than one way. Consider first the identification of an equation of the second sector. Estimation in the way proposed can only be appropriate if (as was the case with the equations of the first sector) an equation of the second sector which is identifiable considering that sector in isolation, is automatically identifiable when considered as part of the full model. Yet this cannot be the case unless more is said, because there just won't be enough restrictions on such equations to do the job. If the number of equations in the k^{th} sector is m_k , with $\sum_{k=1}^C m_k = M$, then the Rank and Order Conditions for identifiability imply that identification when considering the second sector separately requires only $m_k - 1$ restrictions (of the usual linear and homogeneous type). When considering that equation in the context of the full model, we can add

other $(M - m_1 - m_2)$ restrictions corresponding to the fact that $y^3, \dots, y^C$ do not appear in the equation in question (or in any equation of the first two sectors), but this will only give a total of $(M - m_1 - 1)$ restrictions which falls short of the total of $M - 1$ required by the Order Condition. Something is wrong.

Another way to see the problem is to consider the argument made above as to how the block-triangularity of B identifies the equations of the first sector with respect to the equations of higher-numbered sectors. Applying that same argument to the equations of the second sector works all right but only achieves identification with respect to the equations of sectors $3, \dots, C$. This reflects the fact that the m_1 equations of the first sector obey all the same restrictions which the triangularity of B places on the equations of the second sector.

The source of the difficulty becomes apparent if we go back and examine more closely the argument that we can safely treat y^1 as predetermined in estimating the equations of the second sector. It is not true that y^1 depends only on z ; y^1 depends also on u^1 , the disturbances in the first sector's equations. Treating y^1 as predetermined in estimating the equations of the second sector requires that y^1 be uncorrelated (at least in the probability limit) with u^2 , the disturbances in the second sector's equations. If u^1 and u^2 are themselves uncorrelated, this will be the case; if u^1 and u^2 are correlated, however, then so will be y^1 and u^2 . In that case, y^1 cannot be treated as predetermined in the way described, and the second sector cannot be consistently estimated without some explicit consideration of the first sector even though there are no feedback loops connecting them.

As a formal matter, we took care of this in the definition of block-recursive systems by assuming Σ to be block-diagonal (and by not drawing any arrows directly connecting u^1 and u^2 in Figure 3.1). This is enough to give the desired result enabling us to treat the second sector (and similarly all higher-numbered sectors) separately in the way described.

In practice, however, this aspect is often conveniently overlooked, so it is well to realize just how crucial it is. It turns out that zero correlation between u^1 and u^2 (more generally, the block-diagonality of Σ) is just enough to do the job. This is so in the precise sense that if an equation of the second sector is just identified when considering that sector in isolation, then it remains just identified in the system as a whole if u^1 and u^2 are uncorrelated. In that situation, failure of the block-diagonality assumption implies loss of identification for such an equation. Nearly as important for our purposes, even if an equation of the second sector is over identified when considering that sector in isolation (unless u^1 and u^2 are uncorrelated) that equation may cease to be identified in the context of the full model. Moreover, whether it is or not can only be determined by explicitly considering the equations of the first two sectors together and, in any case, the equations of the second sector cannot be consistently estimated by treating that sector separately and taking the variables in y^1 as predetermined.¹ Similar remarks apply to higher-numbered sectors.

Sectoral investigators have got to pay attention to this. The dis-

¹Although this doesn't mean that the equations of the first sector must themselves be estimated. See below.

turbances of a real simultaneous equation model are not merely mathematical abstractions. They represent real political-economic-social forces left out of the model. If such forces cross sectoral lines, then the easy treatment of separate sectors suggested by the block-triangularity of B will break down. Every investigator must give thought to the forces generating the variables he wishes to take as predetermined. He must not only determine that there are no explicit feedback loops preventing him from so taking them but also must consider, so far as possible, the nature and source of the disturbances involved. This is not easy to do, for disturbances by their nature represent influences about which he is likely to know relatively little, but the seductive convenience of assuming no cross-sectoral correlation does not excuse him from making the effort. Faith may be "the evidence of things not seen," but the faith required here is that the "things not seen" are limited in their influence.

Quite a similar situation applies when we consider the case of lagged endogenous variables. Here, of course, there can be no simultaneous feedback effect from current endogenous variables, so that we can think of the lagged endogenous variables as determined in a lower-numbered sector. This corresponds to the block-triangularity of B . What corresponds to the block-diagonality of Σ ?

It is easy to see that the precise assumption required is that the disturbances from the equations determining the lagged endogenous variables be uncorrelated with the disturbances from the equations in which the lagged endogenous variables appear on the right-hand side. If this is so, then those lagged endogenous variables can be treated as predetermined,

but not otherwise unless more is assumed.

This is worth exploring a little further. If a lagged endogenous variable appears on the right-hand side of an equation explaining the current value of the same variable, then the assumption in question is the familiar (but often erroneous) one that the disturbance from that equation not be serially correlated. More generally, however, such lack of serial correlation in each disturbance taken separately is not sufficient (and it may not even be necessary). This is because the lagged disturbance from one equation can be correlated with the current disturbance from another equation even if neither are themselves serially correlated. In general, however, since we would expect "own" serial correlation effects to be the strongest, it would be odd if we were willing to assume them absent without assuming such "cross" effects absent also. It is clearly sufficient that every element of the entire current disturbance vector be uncorrelated with every element of the entire lagged disturbance vector.¹

This discussion of lagged endogenous variables as a special case, however, raises a question with wider implications for our treatment of the block-recursive case. There are well-known ways for handling the case of serially correlated disturbances. In single equations, these allow one in effect to treat the lagged dependent variable as predetermined, and such methods also generalize to simultaneous equation systems.² Doesn't

¹Some discussion of how the fact that within-sector serial correlation is likely to be more important than between-sector serial correlation can be used in sectoral estimation is found in Fisher (1965).

²Fair (1972) is the simplest example.

this mean that lack of block-diagonality of Σ is not necessary in order to take endogenous variables from lower-numbered sectors as predetermined?

The formal answer to this question is that block-diagonality of Σ is indeed not necessary in the presence of other information but is likely to be necessary in practice. The methods referred to for treating serial correlation are methods which assume a particular parametric form for such correlation (usually a first-order form). Provided that assumption is correct, such methods will indeed often permit the treatment of lagged endogenous variables as predetermined; if serial correlation takes a different form, such treatment will lead to inconsistency.

The same thing is true in the general case. If we knew some parametric form taken by the departure of Σ from block-diagonality, then it would often be possible to use that information and go on treating endogenous variables from lower-numbered sectors as predetermined. (It would not always be possible because such information, together with other restrictions, would have to be sufficient to identify the equations of the sector being estimated. This is also true in the serial correlation case¹.) This is not a matter of any practical importance, however, except in the case of lagged endogenous variables. This is because there is generally no earthly reason to suppose that the $(m_j \times m_k)$ correlations between the disturbances of the j^{th} and k^{th} sectors can be expressed in terms of any fewer than $(m_j \times m_k)$ parameters. Hence, there is not generally going to be any trick which will enable us to treat the endogenous variables in question as predetermined unless Σ can be assumed to be block-diagonal.

¹In that case, much of what is involved is the presence of truly exogenous variables. Cf. Fisher (1966, Ch. 6).

Of course, this does not mean that failure of block-diagonality must prevent any consistent estimation of the equations of a sector. (Indeed, even failure of the block-triangularity of B need not prevent such estimation.) What it does mean is that such estimation (assuming identification to be present) has to pay attention to the equations of lower-numbered sectors.

How is this to be done? The basic answer, of course, is instrumental variables. If the endogenous variables from lower-numbered sectors are not to be treated as predetermined, then they must be treated as endogenous. Instrumental variable estimation does this by, in effect, replacing each such endogenous variable by its predicted values from a regression on truly predetermined variables. If this is to be done efficiently the instruments involved must be intelligently chosen. This means that consideration has to be given to the processes generating the endogenous variables in question, thus considering, at least as the major starting point, the equations of the lower-numbered sector from which they come.

Thus, failure of block-diagonality prevents us from simply estimating the equations of a given sector as if they were alone in the world. Note, however, that this does not mean that sectoral treatment becomes impossible and that the full underlying model must be estimated. Not at all. For instrumental variable estimation of the type described, it suffices to know what some (or, preferably, all) of the pre-determined variables are that influence the endogenous variables from lower-numbered sectors. One can know this without a complete specification of the equations describing the inner workings of those sectors (information about a sort of quasi-reduced form is enough); further, one can certainly use such

information without ever having to estimate such equations explicitly, or even, sometimes, having data on all the variables which appear in them.¹

It is for this reason that I have not included block-diagonality in defining what is meant by a "sector", even though it forms a crucial part of the definition of block-recursive systems. Its presence or absence makes a big difference, however, and it is too often assumed present with no justification whatever.²

4. Ignoring Feedbacks in Sectoral Estimation

So far, this rather lengthy discussion of block-recursive systems has supposed that B really is block-triangular. But what if this is not so? We said in defining a sector that what was required was that the feedback loops being ignored not be important, not that they be entirely absent. Clearly, if complete block-triangularity were required for any sectoral estimation, then sectoral estimation would be considerably rarer than it is.

¹Of course, there may be advantages to estimating such equations explicitly. In particular, once they are estimated, estimates of the off-diagonal blocks of Σ can be obtained and used in three-stage least squares or similar methods. Discussion of this departs from our main topic of sectoral estimation, however.

²A special case of this occurs when each sector consists of a single equation and block-diagonality (in this case diagonality) would imply recursiveness and the consistency of ordinary least squares. Econometricians, scenting the home stable, too often gallop over this point.

What then are the consequences for estimation of ignoring small but non-zero feedback effects? In the first place, the answer depends on just how those effects are ignored. At the end of the preceding section we observed that departures from the block-diagonality of Σ could be handled using instrumental variables without explicit estimation of equations not in the sector. The same thing is true of departures from the block-triangularity of B . Even if feedback effects are large, one can estimate one or more equations in a simultaneous model by such methods; this, in a way, is what limited-information methods are all about, allowing one to estimate one equation without using (or, to a large extent, knowing) the prior restrictions and exact nature of the others.

Such a technique is obviously good practice. If one really thinks that feedback effects are large enough to make it dangerous to take certain variables determined outside the sector as predetermined, then it is wise to treat them as endogenous. This can usually be done without a great deal of extra effort merely by specifying some truly predetermined variables which are likely to influence them. The departure from convenience involved in so doing is likely to pay off in estimation consistency.

On the other hand, if feedback effects really are small, then the situation is less clear. By treating a non-sector variable as endogenous and substituting for it its predicted value from a regression on a few instruments, one is certainly sacrificing some efficiency in order to gain what may be relatively little in the way of consistency. In other words, the best instrumental variables are those which (asymptotically, not in the sample) are highly correlated with the right-hand side endogenous variables, provided that the instrumental variables are themselves predetermined. If the instrumental variables available are not so highly

correlated with the extra-sector endogenous variable in question, then using that variable itself may be preferable if the departure from exogeneity is not great.

This brings us to the obvious question of how to tell in advance whether feedback effects are large or small. To some extent, this is like asking "how high is up?" All that can be done is to point out what the correct units are, so to speak. The important thing here is to realize that the fundamental issue is not the correlation between the disturbance in the equation being estimated and the extra-sector endogenous variable in question. Despite (or rather because of) the scale-free nature of the correlation coefficient, what matters is not scale free and depends in a fairly complicated way on all the parameters involved in the feedback. What is involved, in general, is the vector of multiple regression coefficients in the regression of the disturbance in the equation to be estimated on all the variables used as regressors in the last stage of two-stage least squares.¹ While it is true that these all go to zero as the feedback effects ignored approach zero,² they are quite complicated functions of the parameters involved.

To see this, consider the simplest possible example with just two equations and two variables. The equation to be estimated is:

$$(4.1) \quad y = \beta x + \epsilon ,$$

¹ This is an application of Theil's specification error theorem. See Theil (1957).

² See Fisher (1961).

where y and x are scalar variables, β is the parameter to be estimated, and ϵ is a disturbance. However, the variables are also related by:

$$(4.2) \quad x = \alpha y + u$$

where α is a parameter and u is a disturbance. Treating x as predetermined amounts to assuming that α is zero.

Now, in this case, treating x as predetermined leads to the use of ordinary least squares to estimate (4.1). Moreover, since there is only one variable on the right-hand side of that equation, it is very easy to exhibit a detailed formula for the inconsistency involved in doing so; in more general cases, the matter would be even more complicated in detail. Letting b be the least-squares estimator of β , it is obvious that:

$$(4.3) \quad \text{Plim } b = \frac{\text{cov}(x, y)}{\text{var } x} = \beta + \frac{\text{cov}(x, \epsilon)}{\text{var } x}.$$

To work out the last term on the right-hand side, observe that, solving (4.1) and (4.2), we obtain:

$$(4.4) \quad x = \frac{u + \alpha \epsilon}{1 - \alpha \beta}.$$

Assuming that u and ϵ are uncorrelated (otherwise the problems involved in the non-block-diagonality of Σ are upon us), this means that the inconsistency involved in treating x as predetermined is:

$$(4.5) \quad \frac{\text{cov}(x, \epsilon)}{\text{var } x} = \frac{\alpha(1 - \alpha\beta)\text{var } \epsilon}{\alpha^2 \text{var } \epsilon + \text{var } u}.$$

This certainly goes to zero as α does and is monotonic in α for α close enough to zero, but that is about all that can be said.¹ Even in this simplest of cases, trying to decide whether the specification error involved in ignoring feedbacks is small enough not to matter is a complicated business.

Despite this, it is important to realize that for α sufficiently small in the above example, or, more generally, for feedback effects sufficiently weak, such effects can be ignored. Although consistency appears to be naturally a dichotomous property, this is misleading.² Although an estimator is either consistent or inconsistent, what matters is the distance which its probability limit is from the true parameter. In the present case, that distance goes to zero as feedback effects get smaller.³

Return to the example of the milk sector used earlier. The fact that the income of dairy farmers is part of total income means that estimation of milk demand taking total income as exogenous gives inconsistent parameter estimates. Whether or not this matters, however, does not turn on whether the income of dairy farmers is correlated with total income, but rather on whether a change in dairy-farmer income produces a large or a small change in the total. In a logarithmic model, for example, such an effect will be small if the milk sector is a small one even if dairy income and total income are highly correlated. In that case, the feedback effect being ignored in dealing with the milk sector is one which may indeed be negligible.

¹It is also true that the inconsistency goes to zero as the variance of ϵ , reflecting the fact that it is a regression coefficient which occurs in (4.3). Note that this is not true if the variance of u also goes to zero. This is intimately related with the identification properties of the model when one disturbance variance goes to zero. For a further discussion, see Fisher (1966, Ch. 3). Note that, since least squares is not a consistent estimator of (4.1), one could not conclude that because (4.1) fit the data well that the variance of ϵ was small.

²For a fuller discussion, see Fisher (1961).

³The same is true of identifiability. See Fisher (1966, Ch. 3).

Some words of warning concerning ignoring such effects must be given, however. It is wrong to suppose that just because the dairy sector is small, the percentage effect of ignoring the feedback effect in estimating the income elasticity of demand will be small. To see this¹, consider the following.

It is obvious that treating income as exogenous in estimating the consumption function results in substantial inconsistency. Suppose now that we disaggregate total consumption into many parts, one of which is the demand for milk. If we treated income as exogenous in estimating the marginal propensity to consume each good, then, because of the linearity of all relevant estimators, the bias and inconsistency in each of the parts would have to add up to the bias and inconsistency we would get if we treated consumption as a single unit. There is thus no reason to suppose that the percentage effect in estimating the effects of income on the demand for milk will be any less than the percentage effects in estimating the consumption function as a whole.

Put it another way. The basic reason for believing that one can treat income as exogenous in dealing with the demand for milk in the situation we've been exploring is that milk is a small part of the economy. In effect, this says in part that the marginal propensity to consume milk is small. Income gets spent largely on other things. Thus the reason for believing the bias and inconsistency involved to be small is essentially the same reason that prompts us to believe that the effect being measured is also small. No proposition about neglecting effects in a

¹For a more complete discussion, see Christ (1966, pp. 158-60) relating the famous shoelace example given in Koopmans (1949).

percentage sense can be produced this way. You cannot get rid of a large bias simply by hiding its pieces in different places.

On the other hand, this does not mean that one can not treat income as exogenous in analyzing the milk sector (even though it is obviously advisable to avoid this if possible). This is because it's not at all obvious that interest really centers on the percentage errors made in estimating the income effects of the demand for milk. No similar argument applies to the estimation of price effects in the demand equation, for example. There it is likely to be the absolute, rather than the relative, size of the feedback effects being neglected which is important.

But watch out. It is sometimes thought in practice that if an endogenous

variable is treated as exogenous in this way, then the effects of the specification error committed are restricted to the coefficient of that variable in the sectoral equation being estimated. This is not true. In the general case in which that variable is not orthogonal to the other variables in that equation, the effects of such specification error, small or large, spread to the estimates of all the other coefficients in that equation. Moreover, it is not generally true that the largest such effect occurs on the estimate of the coefficient of the directly affected variable.¹ Hence, while the consequences of ignoring a feedback effect in estimation all go to zero with the feedback effect ignored, it is well to be aware that they do tend to spread. In the dairy example, for instance, it will not merely be the estimated income elasticity of demand for milk which will be affected; it is entirely possible that the effects on the estimated price elasticity will be even greater.

It should be noted, however, that if a limited-information type technique is used (two-stage least squares, for example), then the effects of such specification errors will be limited to those equations where the endogenous variable treated as exogenous actually appears provided that such variable is not itself used as an instrumental variable in estimating other equations.² This "quarantining" of specification error is a powerful reason for using such limited-information errors rather

¹For a fuller discussion, see Fisher (1977).

²See Fisher (1961).

than full-information estimators such as three-stage least squares. In the present context, it is also a powerful reason for not using such variables as instrumental variables in estimating equations in which they do not appear unless forced to do so.

In terms of the milk example, it may be necessary and even desirable to treat total income as exogenous in estimating the demand equation in which it appears. It will generally be less obviously desirable that one also wishes to risk messing up the estimates of the supply equation as well. One should do so only if the variable in question (total income) is so important a part of the determinants of the right-hand side endogenous variables in the equations in which it does not appear (the supply equation), that to leave it out would seriously reduce the extent to which such right-hand side variables can be moved about through movements of the remaining instruments. This is likely to be true in the milk example. (The limiting case of this, of course, is where failure to use the variable in question would lose identifiability.¹) Even in this case, it may be entirely possible adequately to substitute for the use of the variable in question by using its predicted value from a first-stage regression on other instruments which determine it, or to use such instruments directly.² As in the case in which the variable in question appears explicitly in the equation being estimated, a trade-off between consistency and efficiency is involved here and it does not seem possible to give more than rather general guidance.

¹ Loss of identifiability can occur in such a case because ignoring the presence of such a variable in sectoral equations other than the one being estimated is tantamount to ignoring an identification restriction, namely, the fact that the variable does not appear in the equation at hand.

² One might wish to do the first if the number of observations are small. This is a sort of truncated Brundy-Jorgenson (1971) device.

5. Simulation and Forecasting: No Feedbacks Ignored

Having estimated the equations of a sectoral model, the investigator now proceeds to exercise them in simulation experiments or in actual forecasting. In so doing, he must continue to consider the fact that his model is embedded in a larger one, for the variables of that larger model appear as exogenous variables in his equations. If, in creating his sector, he has ignored no feedback loops, then he has the relatively simple problem of obtaining time series for those exogenously determined variables. If, as is more likely, he has ignored some small feedback loops (and certainly if he has ignored large ones), then his problem is more difficult because the timepaths of such "exogenous" variables will themselves be affected by the action of his model. In this respect, there is really little difference between simultaneous and lagged feedback effects (except for very short-run forecasting), and we return to the supposition that the ignored effects are of either kind.

Since the problems which arise when there are no feedback effects ignored are interesting in themselves and since they continue to arise in the more general case, I shall discuss them first.

a. Forecasting Exogenous Variables

The first such problem is that of obtaining time series on the

exogenous variables to use for simulation and forecasting. Obviously, this is not a problem if one is going to examine behavior during the sample period. . Moreover, sample period time series for exogenous variables can be used in more interesting ways in conducting simulation experiments examining what the model would have done had exogenous variables behaved differently. Nevertheless, one will generally want to exercise the model to make forecasts or to consider variations in behavior for different future scenarios. For some such exercises, at least, it will be necessary or desirable to use forecasts of the exogenous variables.

If one is lucky, such forecasts will be readily available as outputs of other models, particularly macro models. More generally, however, this will not be the case, following the Iron Law of Econometrics that data are never available in the form in which one wishes to use them. Rather, it is likely that macro models deliver forecasts of variables which are closely related to but not identical with those which one wants to use. One example would be the appearance in a sectoral model of the output of non-ferrous metal industries where a macro model predicts all industrial output. Another would be the prediction by a macro model of demand for consumer durables where what is really needed in the sectoral model is demand for automobiles.

In such a situation, some linkage has got to be provided between the actually forecasted variables and those which the investigator would like to have. Obviously, the preferable way to do this in some overall sense would be to construct a structural model which links the sectoral model to the macro model used. Unfortunately, this is likely to prove as great a task as the original construction of the sectoral model of

primary interest, if not a greater one. Moreover, the pay-off from doing it is not likely to be very great in terms of the objectives of the sectoral model builder (although, it can obviously be a valuable aid to understanding the more general economy). This is because the primary aim here is to obtain reasonable forecasts for exogenous variables on which to build varying simulation and forecasting experiments with the sectoral model. While it is obviously desirable that such forecasts should be as good as possible, fine tuning in this regard may not be worth the vast effort that may be involved, particularly if there are other exogenous variables which are not themselves closely related to available macro-model outputs (prices of other commodities, for example), and hence must be forecast by less rigorous means.

Moreover, like it or not, econometricians should admit that the principal value of structural macro models has been in the understanding which they impart as to how the economy works, and not primarily in the improved precision which they lend to forecasting. Detailed, structural adaptation of macro-model forecasts to produce forecasts of the particular exogenous variables which appear in the model is unlikely to do particularly better than a relatively simpler procedure. That procedure is to take sample period data and regress, using ordinary least squares, the exogenous variables to be predicted on the most closely related variables which appear in the macro model, using the estimated regressions to generate the desired forecasts from the forecasts of the macro model.

One might ask, however, why not go the whole way and generate forecasts from naive models without using any structural macro model at all? Obviously, if one thinks that the macro model one is using produces fore-

casts which are actually inferior to those of a naive model, one ought to discard it. The comment made above as to the added precision of forecasts achieved through structural macro models should not be taken to mean that such models provide no aid to forecasting at all; it is merely that they do so at a heavy price in time, effort, and money. By piggy-backing on the forecasts of an available macro-model, the sectoral investigator can obtain some of the fruits of such inputs. This can be potentially very valuable, particularly since structural forecasts are likely to prove superior to naive ones at precisely those moments when forecasting is important, when things stop going up and down together and turning points occur.¹ I favor using as much structural information as reasonably possible; indeed, before constructing the regressions referred to, the investigator ought to give considerable thought as to what variables the structure of the economy suggests ought to appear in them. Much more than this, however, is not likely to be worth the cost.

A different question naturally arises, however. If, at the end of the estimation process, one is going to be forced to use variables in a particular form because those are the variables which are forecast by an available macro model, ought one not to build ones model to take this into account in the first place, and use such variables directly rather than use variables for which one has later to generate relatively ad hoc forecasts? There is something in this, but, I think, not very much.

Obviously, there is no point in producing complications for their own sake. If there is no particular reason for using an inconvenient variable rather than a conveniently forecast one, then one ought to use the latter.

¹Lesnoy (1958) provides some experimental verification of this. See Fisher (1965).

Where there is such a reason, however, I think it would be a mistake to substitute in order to avoid the forecasting problem at the end. The purpose of sectoral modelling (or, indeed, of all structural modelling) is not merely the generation of easy forecasts. Rather, it includes the discovery of the mechanisms operating in the economy. Using the wrong variables because they are convenient can only hinder this aim.

Put it another way. The sectoral investigator generally has to experiment with the specification of his model, interacting with the data and testing (perhaps informally) various hypotheses as he goes along. Using the wrong variables will cloud his vision. He will never be quite sure whether the failure of some trial equation is due to underlying, actual effects or to his having messed things up for later convenience. He can use the right variables and find out, but, if the equation then seems successful he will face the same forecasting problems as before.

More formally, substituting convenient variables for correct ones involves, in effect, the implicit substitution into the sectoral model's structural equations of equations relating correct variables to convenient ones. These latter, auxiliary equations are those which would have to be estimated directly were the correct variables used in the sectoral model; in their simplest form, they are the regression equations described above. Since such equations are unlikely to be rich in structural detail and since the investigator is properly unlikely to be willing to expend large amounts of effort on them, why should they be used in such a way as to contaminate the sectoral modelling process itself? Surely it is better to treat them explicitly as what they are and to estimate them directly. If they fit the data well, then there is little inconvenience

involved in going from macro-model to exogenous-variable forecasts. If they don't, then the justification for substituting convenient for correct variables must be weak. It seems definitely better to know rather than to guess at this.

A somewhat more appealing way of doing the same thing presents similar problems. Since the auxiliary equations involved will be used together with those of the sectoral model to generate forecasts, and since those forecasts will have to go to the endogenous variables given the macro-model outputs, why not estimate such auxiliary equations together with the equations of the model so as to make the forecasts as good as possible. For example, why not use a full-information, maximum-likelihood technique on all of them together? The trouble here is much the same as before. The auxiliary equations are less soundly based than are the equations of the model; full-information techniques tend to carry specification errors from one equation to another. It does not seem desirable to contaminate the estimation of the equations of the sector because of problems with the forecasting equations. To put it over-simply, one would hardly want to change ones estimates of the structural parameters because the auxiliary equations hadn't fit well.

It is conceivable, however, that estimation techniques could be devised which would allow the auxiliary equations to be estimated bringing to bear the information in the structural equations but which would leave the structural estimates invariant with respect to matters affecting the auxiliary equations. This would be a more interesting possibility, but, so far as I know, it remains to be worked out. Further, it has so much the flavor of getting something for nothing that general pessimism suggests

that it is not possible. In any case, barring this possibility, it seems best to treat auxiliary equations separately.

b. Correlated Disturbances

A somewhat different problem in simulation and forecasting arises not just in sectoral models, but in modelling generally. Disturbances tend to be serially correlated. We have already discussed some of the consequences of this for estimation; there are also consequences for simulation and forecasting.

In its simplest form, the problem comes down to this: One is forecasting one period ahead. For the last, non-forecast period, the predicted values of the endogenous variables will generally not be the same as their actual values. The differences (residuals) are estimates of the disturbance values for that period. Given that disturbances tend to be serially correlated, those estimates provide some information about the values of the disturbances in the first forecast period. Should one use that information in generating a forecast? Simply put, if a particular disturbance is positively serially correlated and one has estimated the correlation coefficients involved, then there is a predicted value for that disturbance one period later, given the value of the residual in the last non-forecast period. Should one adjust the value which the model would otherwise predict for a particular endogenous variable by adding in the predicted value of the disturbance?

Moreover, this question arises in somewhat more general context. For one thing, there is no need to restrict ones attention to forecasting a single period ahead. Information on the value of a disturbance in the

last non-forecast period provides information on the value of that disturbance two or three, or more, periods ahead. Of course, that information becomes less reliable the farther ahead one wants to go, but it is there nonetheless.

Related to this, and more directly related to our general concern of the relations between sectoral and macro models, is the fact that disturbances from different equations in the model tend to be correlated and, more particularly, that the disturbances from the equations generating the exogenous variables in the sectoral model may be correlated with the disturbances of the sectoral model itself. This non-block diagonality of Σ was discussed above in terms of its implications for estimation. In terms of forecasting another question arises as to whether one ought to use information on such correlation. More specifically, the equations of the macro model (perhaps together with auxiliary equations of the type discussed earlier in this section) generate predicted values for exogenous variables. By examining the difference between those values and the actual values of the exogenous variables for the last non-forecast period, one obtains information on the disturbances in those equations. If those disturbances are auto-correlated this fact can be used to generate predicted disturbances for the first (or later) forecast period. Since those disturbances are correlated with the disturbances from the sectoral model of interest (non-block diagonality of Σ), one can go on to generate predictions of disturbances in the sectoral model itself. Should this procedure be used in generating forecasts?

There is something to be said on both sides of these issues. In favor of not using such methods is, of course, consideration of their somewhat

roundabout nature, a consideration which applies with increasing force the further in the future one wishes to forecast; however, some account can probably be taken of this in terms of generating errors of forecast. More important, estimates of the correlation structure of the errors do not rest upon structural information at all. Disturbances are the things we can't (or at least don't) explain. We put in correlations between them to allow for generality and as a substitute for analysis of the real forces which make them up. By so doing, we can often improve the quality of our estimates in more than one way. Yet, to rely on such correlations to aid us in forecasting is to rely somewhat blindly on the continuation of relations which we simply do not understand. This may be a dangerous thing to do.

On the other hand, to the extent that we are forecasting relatively close to the sample period, there may be no reason to suppose that a correlation structure which well explains what's been going on in the sample period cannot be expected to continue explaining it for a few periods yet to come. Since forecasting is a tough business, it may be important to use all the help one can get and one ought not to throw away information, even information on the distribution of the disturbances, unless there is some serious reason for supposing that it's wrong.

There is something to be said for both of these views. My own experience (somewhat limited) is that one ought to use the information. Simulation experiments with the world copper model described in Fisher, Cootner and Baily (1972) seemed to end up a bit more sensibly when one paid attention to serial correlation than when one didn't. However, it is hard to be sure. In practice, it may be sensible to test which way

one ought to go by trying them out in forecasting some periods after the sample period for which data are available. Since it is generally true that by the time one is ready to do this data beyond the sample period have become available, this need not involve a wait.

6. Simulation and Forecasting: Ignoring Feedbacks

When there are feedback loops connecting the sectoral model to the underlying model which have been ignored in constructing the former, all the problems just discussed continue to exist, but there are other ones as well. There are two main cases to consider here. The first such case is one in which one has available the outputs of an existing macro model and the feedbacks which have been ignored are those connecting the sectoral model to the macro model. The second case is that in which no convenient macro model exists or in which the feedbacks involved cannot sensibly be thought of as relating only to macro variables. It will turn out that, in some logical sense, it is the second case that presents all the difficulties.

Consider, then, the case in which there is a complete macro model in existence and the feedbacks being ignored are feedbacks that connect the sectoral model to the macro model. In terms of the milk market example used several times above, this can happen because there is a model of the determination of national aggregates, the variables from which have been used as exogenous variables in the model of the milk market,

and the feedback being ignored is that between the income received by dairy farmers and total income.

In some real sense there is no problem here, or, at least, the problem is not at the level of the sectoral model. In this example, if the macro model is a good one, then all the influences determining total income are included in it, although in a very aggregative way. The fact that total income affects the demand for milk, which in turn affects the income of dairy farmers, is not different from the fact that total income affects the demand for any commodity, which in turn affects the income of the producers of that commodity. If such effects are so important as to require explicit representation in order to generate a satisfactory model of the determination of total income, then such effects should have been modelled in the macro model in the first place. If one considers that such effects do not have to be detailed in order to get an adequate explanation of total income, but instead one is content with representing such feedback loops in general terms (for instance, through the consumption function, as representing aggregate demand), then, in using the macro model to generate income forecasts for use in forecasting and simulation with a sectoral model of the milk market, one need not pay attention to feedback loops of the sort described. If one doesn't think that the macro model adequately accounts for such effects (and this is certainly possible), then one is in the case in which one doesn't have a completely adequate model to start with. This is the second case listed above, and the one to which I shall now turn.

One possibility in this case is to try to build an explicit representation of the feedback loops involved. In effect, this means describing

how the outputs of the sectoral model affect the inputs of the macro model. If this is well done, then one is building a part of the macro model oneself; indeed, the problems involved in doing this are much the same as those faced by a macro-model builder who decentralizes his enterprises and allows others to build sectoral models for inclusion in the overall macro model. Unless this is actually the way in which the problem arises, however, the sectoral-model builder is unlikely to find that the outputs for his model exactly (or even closely) match the required inputs for the macro model. If he is to attempt to match them, then some more work needs to be done.

In essence, what is involved here is the exact counterpart of the situation discussed in the previous section in which the output for the macro model had to be adapted for forecasting purposes in predicting the exogenous variables of the sectoral model. Once again, the simplest thing to do is likely to be the construction of auxiliary equations, this time regressing the variables in the macro model likely to be affected by sectoral model outputs on those outputs. It should not, of course, come as a surprise that one runs into the same problem turned around, because (see Figure 2.1 above) in this case, the only way one can tell which is the sectoral model and which is the macro model is in terms of labels.

As a practical matter, however, it is not likely frequently to be the case that the problems of ignored feedback loops can be even approximately overcome by the simple expedient of grafting in the manner described. The more common situation is likely to be one in which one has reason to believe that the feedback loops are not negligible but the

macro-model availability does not lend itself readily to incorporating them. Alternatively, the ignored feedback loops do not involve a macro model at all, but rather involve variables from other sectors which may be predetermined (or approximately so) when doing estimations, but cannot really be taken to be exogenous in constructing forecasts.

An example of the latter sort of problem can be found in Fisher, Cootner, and Baily (1972). In that model of the world copper market, the price of aluminum appears in some of the demand equations. Because it was unnecessary to do so for estimation purposes, no explicit attention was given to the influence of the price of copper on the price of aluminum.¹ In forecasting the price and demand for copper, however, attention has to be given to that sort of feedback. Obviously, one can say that the failure to estimate equations involving the aluminum price simply shows that the model was not one of a complete sector; but, had one explicitly estimated such equations, not only would the work have increased substantially, but also the equations involved would have introduced still further variables (the price of steel for example), and the same problem would be encountered at a new level. If sectoral modelling is to mean anything at all in practice, one is going to run into this kind of problem.

Unfortunately, there isn't a very satisfactory solution to the problem. What one has to do, I think, is to make varying assumptions about the nature and strengths of the feedbacks involved and then do forecasting and simulation experiments on those different assumptions. Since the feedbacks being ignored ought to be small (if they are large, then one really hasn't got hold of a complete sector in any sense), changes in the assumptions within a plausible range may not make too much difference.

¹ Aluminum price entered the model only lagged, and allowance for serial correlation was made. Note that even had it entered without a lag, use of instrumental variables would have permitted estimation without very explicit consideration of the structure involved (see above).

The difficulties here are obviously worst as involve forecasting. For simulation, they are not so severe. Simulation experiments have several objects. A frequent one is to examine the difference which it makes if certain parameters are changed or if certain exogenous variables take on different time paths. In a sense, doing simulation experiments by varying the parameters describing small but ignored feedback loops is no different from doing simulation experiments with other parameters. The principal difference, of course, is that one is guessing the parameter values with which one starts rather than estimating them. This is a common affair for promulgators of the type of dynamic modelling practiced by J. Forrester,¹ but it isn't good practice and should be avoided in econometric models whenever possible, especially for important parameters of the model. Presumably, the parameters here under discussion are not very important.

Simulation experiments can be carried on with guessed parameters for feedback loops under discussion in another way as well. It is likely that those parameters do affect the time paths generated for model predictions. Nevertheless, it is entirely possible that the difference between two such time paths generated by varying exogenous variables or policy parameters in the model is less affected by the exact size of the feedback parameters. If this is so, then the effects of changes in exogenous variables or of changes in policy parameters can be simulated more or less accurately, simply by taking a plausible guess as to feedback parameters and holding it constant while other simulation experiments are performed. The easiest thing to do, indeed, is to perform the simulation experiment ignoring feedbacks altogether -- this amounts

¹For example (1961), (1969), and (1971).

to setting the parameters involved at zero. This is unlikely to give good forecasts if the feedbacks are really non-negligible, but it may very well tell one how good forecasts would change as a result of variations in the effects in which one is really interested. Obviously, it is sensible to check the view that such differences can adequately be explored in this matter by trying out some variations on the feedback parameters involved.

Simulation experiments performed in this way can be very revealing. One ought always to remember, I think, that econometric models of any sort are aids to the understanding rather than mechanical devices for forecasting. By exercising the models in simulation experiments, one can gain understanding into the basic mechanism of what is going on. Moreover, as time progresses it will be possible to check that understanding. In practice, by the time one has a model set up and is ready to go, data will have become available for a few more periods. The problem involved in neglecting feedback loops of the sort described is, after all, one of accepting (or making) outside forecasts of variables treated as exogenous in the sectoral model while ignoring the effects of the model on the variables thus being forecast. An obvious thing to do is to see whether the forecast of the endogenous variables predicted by the model itself would be materially improved if the actual values of the nearly exogenous variables involved in the feedback loops were used rather than initial forecasts of those variables.¹ Even where no post-sample period data are available, one can make experiments like this in trying to forecast for part or all of the sample period. If the model

¹ Note the relation to the problem, discussed, above of whether or not to use information on the disturbances. Here, it is the disturbances of the macro-model which are involved.

forecasts its own endogenous variables well in both cases, or if its forecasts are materially improved by the use of actual data on exogenous variables rather than predicted data, then one has a fairly good indication that the understandings gained with the model are real ones and that further effort ought to be put into forecasting the exogenous input.

Obviously, this can be true whether or not there are feedback loops which have been ignored. If such loops are relatively unimportant, then improving such forecasts will be somebody else's job. If they are not unimportant, then it may still be somebody else's job, but he or she may be grateful for the help your model has given.

References

Brundy, J.M. and D. W. Jorgenson (1971). "Efficient Estimation of Simultaneous Equations by Instrumental Variables," Review of Economics and Statistics 53 (August), pp. 207-224.

Christ, C. (1966). Econometric Models and Methods. (New York: John Wiley & Sons).

Fair, R. (1970). "The Estimation of Simultaneous Equation Estimation," Econometrica 38 (May), pp. 507-516.

Fisher, F.M. (1961). "On the Cost of Approximate Specification in Simultaneous Equation Estimation," Econometrica 29 (April), pp. 139-170. Reprinted in A. Ando, F. Fisher, and H.A. Simon, eds., Essays on the Structure of Social Science Models, (Cambridge, Mass.: M.I.T. Press, 1963), pp. 32-63.

_____. (1965). "Dynamic Structure and Estimation of Economy-Wide Econometric Models," (a) Ch. 15 in The Brookings Quarterly Econometric Model of the United States, J. Duesenberry, G. Fromm, L. Klein, and E. Kuh, eds., (Chicago: Rand-McNally Publishing Co. and Amsterdam: North-Holland); and (b) Proceedings of the Study Week (October 1963) on "Le Rôle de l'Analyse Econometrique dans la Formulation de Plans de Développement," Pontifical Academy of Sciences, Vatican City, pp. 385-447.

Fisher, F.M. (1966). The Identification Problem in Econometrics,
(New York: McGraw-Hill). Reprinted by Robert E.
Krieger Publishing Co., Inc., 1976).

_____ (1970a). "A Correspondence Principle for Simultaneous
Equation Estimation, Econometrica 38 (January), pp. 73-92.

_____ (1970b). "Simultaneous Equation Estimation: The State
of the Art," IDA Economic Papers (July).

_____ (1977). "The Effect of Simple Specification Error on the
Coefficients of 'Unaffected' Variables," M.I.T. Working
Paper #194 (January). Forthcoming in a volume in memory
of Ta-Chung Liu, edited by T.S. Tsiang, L. Klein and M.
Nerlove.

Fisher, F.M. and P. Cootner, in association with M. Baily (1972). "An
Econometric Model of the World Copper Industry," The Bell
Journal of Economics and Management Science 3 (Autumn),
pp. 568-609.

Forrester, J. (1961) Industrial Dynamics.

(1969) Urban Dynamics.

(1971) World Dynamics. (revised 2nd ed., 1973).

(Cambridge, Mass.: MIT Press).

Koopmans, T.C. (1949). "The Bias in Single Equation Methods of Estimating
Behavior Equations Relating to a Small Sector of the

Economy," Cowles Discussion Paper, Statistics, #336 (mimeo).

Lesnoy, S. (1958). "Limited Information versus Reduced From Least Squares in Prediction." Unpublished; (presented at Cambridge meeting of Econometric Society, 1958).

Theil, H. (1957). "Specification Errors and the Estimation of Economic Relationships," Review of the International Statistical Institute 25, pp. 41-51.

Wold, H., in association with L. Jeréen. (1953). Demand Analysis. (New York: John Wiley & Sons).

SEP 30 2008

[REDACTED]

[REDACTED]

MIT LIBRARIES



3 9080 004 415 433

MIT LIBRARIES



3 9080 004 415 441

MIT LIBRARIES



3 9080 004 446 370

MIT LIBRARIES



3 9080 004 446 388

MIT LIBRARIES



3 9080 004 446 396

MIT LIBRARIES



3 9080 004 446 404

MIT LIBRARIES



3 9080 004 446 412

MIT LIBRARIES



3 9080 004 446 420

MIT LIBRARIES



3 9080 004 446 438

MIT LIBRARIES



3 9080 004 446 446

